Optimization of intractable expectations

Suppse you want to minimize an objective function of the form

$$\overset{*}{\eta} := \underset{\eta}{\operatorname{argmin}}\, \mathbb{E}_{\mathbb{P}(z)}\left[f(z|\eta)\right] := \underset{\eta}{\operatorname{argmin}}\, \ell(\eta),$$

where $\mathbb{P}(z)$ is known, but the expectation is not available in closed form.

Optimization of intractable expectations

Suppse you want to minimize an objective function of the form

$$\overset{*}{\eta} := \underset{\eta}{\mathrm{argmin}}\, \mathbb{E}_{\mathbb{P}(z)}\left[f(z|\eta)\right] := \underset{\eta}{\mathrm{argmin}}\, \ell(\eta),$$

where $\mathbb{P}(z)$ is known, but the expectation is not available in closed form.

When does this happen?

- Stochastic control (e.g. you have a factory, and supply and demand are random)
- Black box variational inference (our interest; we'll define it later)

Optimization of intractable expectations

Suppse you want to minimize an objective function of the form

$$\overset{*}{\eta} := \underset{\eta}{\operatorname{argmin}} \, \mathbb{E}_{\mathbb{P}(z)} \left[f(z|\eta) \right] := \underset{\eta}{\operatorname{argmin}} \, \ell(\eta),$$

where $\mathbb{P}(z)$ is known, but the expectation is not available in closed form.

When does this happen?

- Stochastic control (e.g. you have a factory, and supply and demand are random)
- Black box variational inference (our interest; we'll define it later)

What can you do? There are two options, both using the Monte Carlo (MC) estimate

$$\hat{\ell}(\eta) := \frac{1}{N} \sum_{n=1}^{N} f(z_n|\eta) \approx \ell(\eta).$$

Optimization of intractable expectations

Suppse you want to minimize an objective function of the form

$$\overset{*}{\eta} := \underset{\eta}{\operatorname{argmin}}\, \mathbb{E}_{\mathbb{P}(z)}\left[f(z|\eta)\right] := \underset{\eta}{\operatorname{argmin}}\, \ell(\eta),$$

where $\mathbb{P}(z)$ is known, but the expectation is not available in closed form.

When does this happen?

- Stochastic control (e.g. you have a factory, and supply and demand are random)
- Black box variational inference (our interest; we'll define it later)

What can you do? There are two options, both using the Monte Carlo (MC) estimate

$$\hat{\ell}(\eta) := \frac{1}{N}\sum_{n=1}^{N} f(z_n|\eta) \approx \ell(\eta).$$

- Stochastic gradient (SG)
  - Update with $\eta^{i} = \eta^{i-1} - \rho \nabla_{\eta} \hat{\ell}(\eta)$ for some step size ρ (new z_n every step)
  - Approximately minimizes the exact objective

Optimization of intractable expectations

Suppse you want to minimize an objective function of the form

$$\overset{*}{\eta} := \underset{\eta}{\operatorname{argmin}} \, \mathbb{E}_{\mathbb{P}(z)} \left[f(z|\eta) \right] := \underset{\eta}{\operatorname{argmin}} \, \ell(\eta),$$

where $\mathbb{P}(z)$ is known, but the expectation is not available in closed form.

When does this happen?

- Stochastic control (e.g. you have a factory, and supply and demand are random)
- Black box variational inference (our interest; we'll define it later)

What can you do? There are two options, both using the Monte Carlo (MC) estimate

$$\hat{\ell}(\eta) := \frac{1}{N} \sum_{n=1}^{N} f(z_n|\eta) \approx \ell(\eta).$$

- Stochastic gradient (SG)
    - Update with $\eta^i = \eta^{i-1} - \rho \nabla_\eta \hat{\ell}(\eta)$ for some step size ρ (new z_n every step)
    - Approximately minimizes the exact objective
- Sample average approximation (SAA)
    - Find $\hat{\eta} := \operatorname{argmin}_\eta \hat{\ell}(\eta)$ for fixed z
    - Exactly minimizes approximate objective

Optimization of intractable expectations

Suppse you want to minimize an objective function of the form

$$\overset{*}{\eta} := \underset{\eta}{\operatorname{argmin}} \, \mathbb{E}_{\mathbb{P}(z)} \left[f(z|\eta) \right] := \underset{\eta}{\operatorname{argmin}} \, \ell(\eta),$$

where $\mathbb{P}(z)$ is known, but the expectation is not available in closed form.

When does this happen?

- Stochastic control (e.g. you have a factory, and supply and demand are random)
- Black box variational inference (our interest; we'll define it later)

What can you do? There are two options, both using the Monte Carlo (MC) estimate

$$\hat{\ell}(\eta) := \frac{1}{N} \sum_{n=1}^{N} f(z_n|\eta) \approx \ell(\eta).$$

- Stochastic gradient (SG)
  - Update with $\eta^{i} = \eta^{i-1} - \rho \nabla_{\eta} \hat{\ell}(\eta)$ for some step size ρ (new z_n every step)
  - Approximately minimizes the exact objective
- Sample average approximation (SAA)
  - Find $\hat{\eta} := \operatorname{argmin}_{\eta} \hat{\ell}(\eta)$ for fixed z
  - Exactly minimizes approximate objective

Which is better? **It depends.**